0590
/215

#2

OIPE

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/10/023,929   TIME: 08:42:53

Input Set : N:\Crf3\RULE60\10023929.raw
Output Set: N:\CRF3\02112002\J023929.raw

SEQUENCE LISTING
```
  3 (1) GENERAL INFORMATION:
  5        (i) APPLICANT: Hillman, Jennifer L.
  6                       Goli, Surya K.
  8       (ii) TITLE OF INVENTION: NOVEL HUMAN MLS3 PROTEIN
 11      (iii) NUMBER OF SEQUENCES: 6
 13       (iv) CORRESPONDENCE ADDRESS:
 14            (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 15            (B) STREET: 3174 Porter Drive
 16            (C) CITY: Palo Alto
 17            (D) STATE: CA
 18            (E) COUNTRY: USA
 19            (F) ZIP: 94304
 21        (v) COMPUTER READABLE FORM:
 22            (A) MEDIUM TYPE: Diskette
 23            (B) COMPUTER: IBM Compatible
 24            (C) OPERATING SYSTEM: DOS
 25            (D) SOFTWARE: FastSEQ for Windows Version 2.0
 27       (vi) CURRENT APPLICATION DATA:
 28            (A) APPLICATION NUMBER: US/10/023,929
 29            (B) FILING DATE: 17-Dec-2001
 35            (C) CLASSIFICATION:
 37      (vii) PRIOR APPLICATION DATA:
 33            (A) APPLICATION NUMBER: US/09/205,449
 34            (B) FILING DATE: 1998-12-04
 38            (A) APPLICATION NUMBER: 08/805,965
 39            (B) FILING DATE: 1997-02-25
 41     (viii) ATTORNEY/AGENT INFORMATION:
 42            (A) NAME: Billings, Lucy J.
 43            (B) REGISTRATION NUMBER: 36,749
 44            (C) REFERENCE/DOCKET NUMBER: PF-0223 US
 46       (ix) TELECOMMUNICATION INFORMATION:
 47            (A) TELEPHONE: 415-855-0555
 48            (B) TELEFAX: 415-845-4166
 49            (C) TELEX:
 52 (2) INFORMATION FOR SEQ ID NO: 1:
 54        (i) SEQUENCE CHARACTERISTICS:
 55            (A) LENGTH: 262 amino acids
 56            (B) TYPE: amino acid
 57            (C) STRANDEDNESS: single
 58            (D) TOPOLOGY: linear
 60      (vii) IMMEDIATE SOURCE:
 61            (A) LIBRARY: BRAITUT02
```
C--> 28 (before line 28)
C--> 29 (before line 29)

ENTERED

```
 62           (B) CLONE: 762280
 64     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 66   Met Phe Arg Phe Met Arg Asp Val Glu Pro Glu Asp Pro Met Phe Leu
 67    1               5              10              15
 68   Met Asp Pro Phe Ala Ile His Arg Gln His Met Ser Arg Met Leu Ser
 69              20              25              30
 70   Gly Gly Phe Gly Tyr Ser Pro Phe Leu Ser Ile Thr Asp Gly Asn Met
 71          35              40              45
 72   Pro Gly Thr Arg Ala Ala Ser Arg Arg Met Gln Gln Ala Gly Ala Val
 73       50·             55              60
```
W--> 74   Xaa Pro Phe Gly Xaa Leu Gly Met Ser Gly Gly Phe Met Asp Met Phe
```
 75    65              70              75              80
```
W--> 76   Gly Met Met Asn Asp Met Xaa Gly Asn Met Glu His Met Thr Ala Gly
```
 77              85              90              95
 78   Gly Asn Cys Gln Thr Phe Ser Ser Ser Thr Val Ile Ser Tyr Ser Asn
 79             100             105             110
 80   Thr Gly Asp Gly Ala Pro Lys Val Tyr Gln Glu Thr Ser Glu Met Arg
 81             115             120             125
 82   Ser Ala Pro Gly Gly Ile Arg Glu Thr Arg Arg Thr Val Arg Asp Ser
 83          130             135             140
 84   Asp Ser Gly Leu Glu Gln Met Ser Ile Gly His His Ile Arg Asp Arg
 85    145             150             155             160
 86   Ala His Ile Leu Gln Arg Ser Arg Asn His Arg Thr Gly Asp Gln Glu
 87             165             170             175
 88   Glu Arg Gln Asp Tyr Ile Asn Leu Asp Glu Ser Glu Ala Ala Ala Phe
 89             180             185             190
 90   Asp Asp Glu Trp Arg Arg Glu Thr Ser Arg Phe Arg Gln Gln Arg Pro
 91             195             200             205
 92   Leu Glu Phe Arg Arg Leu Glu Ser Ser Gly Ala Gly Gly Arg Arg Ala
 93          210             215             220
 94   Glu Gly Pro Pro Arg Leu Ala Ile Gln Gly Pro Glu Asp Ser Leu Pro
 95    225             230             235             240
 96   Asp Ser Pro Ala Ala Met Thr Gly Glu Gly Pro Gly Ala Ser Ala Leu
 97             245             250             255
 98   Leu Tyr Arg Leu Arg Gly
 99             260
101 (2)  INFORMATION FOR SEQ ID NO: 2:
103       (i) SEQUENCE CHARACTERISTICS:
104           (A) LENGTH: 1322 base pairs
105           (B) TYPE: nucleic acid
106           (C) STRANDEDNESS: single
107           (D) TOPOLOGY: linear
109      (vii) IMMEDIATE SOURCE:
110           (A) LIBRARY: BRAITUT02
111           (B) CLONE: 762280
113      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
115   GGGGGGCGTA CGGAGGTGGC AGCTGTGGGA GGAGGCGGCG TGGAAGGCCG AGGAGCTCAA    60
116   GCCCGGACCA ATCCCCACGT TCCGGGCCGC CACCCTGACC CTGCAGCGTA CCGGGAAGCG   120
117   AAACCGGCCG GATGGGCCGC TGAGCCCGAA TCGGGCACTG TGTGGAGCCC CCTGGAGCTG   180
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/10/023,929   TIME: 08:42:53

Input Set : N:\Crf3\RULE60\10023929.raw
Output Set: N:\CRF3\02112002\J023929.raw

```
118 AGATCAGGAT GTTCCGCTTC ATGAGGGACG TGGAGCCTGA GGATCCCATG TTCCTGATGG  240
119 ATCCCTTTGC TATTCACCGT CAGCATATGA GCCGTATGTT GTCAGGTGGC TTTGGATATA  300
120 GCCCCTTCCT CAGCATCACA GATGGCAACA TGCCAGGGAC CAGGGCTGCC AGCCGCCGGA  360
121 TGCAGCAGGC TGGAGCTGTC TNCCCCTTTG GGNTGCTGGG AATGTCGGGT GGTTTCATGG  420
122 ACATGTTTGG GATGATGAAT GACATGNTTG GAAACATGGA ACACATGACA GCTGGAGGCA  480
123 ATTGCCAGAC CTTCTCATCT TCCACTGTCA TCTCCTACTC CAATACGGGT GATGGTGCCC  540
124 CCAAGGTCTA CCAAGAGACA TCAGAGATGC GCTCGGCACC AGGCGGGATC CGGGAGACAC  600
125 GGAGGACTGT TCGGGATTCA GACAGTGGAC TGGAGCAGAT GTCCATTGGG CATCACATCC  660
126 GGGACAGGGC TCACATCCTC CAGCGCTCCC GAAACCATCG CACGGGGGAC CAGGAGGAGC  720
127 GGCAGGACTA TATCAACCTG GATGAGAGTG AGGCCGCAGC GTTTGATGAC GAGTGGCGGC  780
128 GGGAGACCTC CCGATTCCGG CAGCAGCGTC CCCTGGAGTT TCGGCGGCTT GAGTCCTCAG  840
129 GGGCTGGGGG ACGAAGGGCG GAGGGGCCTC CCCGCCTGGC CATCCAGGGA CCTGAGGACT  900
130 CCCTTCCCGA CAGTCCCGCC GCTATGACTG GTGAGGGCCC CGGGGCCTCA GCTCTCTTGT  960
131 ACAGGCTGAG AGGCTGAGAA ATCATCCCCT GAATAACTTT TTCCTCTCGA TTCCCATCCC 1020
132 CAATTTAATA TTAAATTAAC AGGCAAGCCG GCCCCCACCT CTCCCTGGGG GTCTCAGGGA 1080
133 GAACCTTTCA CGGCACCCTT TCCCTACCTT TTCCTTCTTT AATCTCCTGG TTTACCATTG 1140
134 ATGACTTCGG CTCTGCATCT ACTTACTTGA TTTTTCATTC TGCCACTTCA TCTTCAAACC 1200
135 CCCTCACCTT TCCCATCCTA CTCCTGCCAT GCATTGAAGG GTCAATGCAT TTTGGGGTGA 1260
136 GNTTNGGTTT AGGGGCCCCT TCATNCCTNA GCTACCTGGG TCTTTGCCCA ACTTTTCTCA 1320
137 GA                                                                1322
```

139 (2) INFORMATION FOR SEQ ID NO: 3:

141   (i) SEQUENCE CHARACTERISTICS:
142   (A) LENGTH: 268 amino acids
143   (B) TYPE: amino acid
144   (C) STRANDEDNESS: single
145   (D) TOPOLOGY: linear
147   (vii) IMMEDIATE SOURCE:
148   (A) LIBRARY: GenBank
149   (B) CLONE: 1066392
151   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
153 Met Phe Arg Met Leu Asn Ser Ser Phe Glu Asp Asp Pro Phe Phe Ser
154   1           5                  10              15
155 Glu Ser Ile Leu Ala His Arg Glu Asn Met Arg Gln Met Ile Arg Ser
156            20                  25                  30
157 Phe Ser Glu Pro Phe Gly Arg Asp Leu Leu Ser Ile Ser Asp Gly Arg
158            35                  40                  45
159 Gly Arg Ala His Asn Arg Arg Gly His Asn Asp Gly Glu Asp Ser Leu
160         50                  55                  60
161 Thr His Thr Asp Val Ser Ser Phe Gln Thr Met Asp Gln Met Val Ser
162 65                  70                  75                  80
163 Asn Met Arg Asn Tyr Met Gln Lys Leu Glu Arg Asn Phe Gly Gln Leu
164                 85                  90                  95
165 Ser Val Asp Pro Asn Gly His Ser Phe Cys Ser Ser Ser Val Met Thr
166            100                 105                 110
167 Tyr Ser Lys Ile Gly Asp Glu Pro Pro Lys Val Phe Gln Ala Ser Thr
168            115                 120                 125
169 Gln Thr Arg Arg Ala Pro Gly Gly Ile Lys Glu Thr Arg Lys Ala Met
170         130                 135                 140
171 Arg Asp Ser Asp Ser Gly Leu Glu Lys Met Ala Ile Gly His His Ile
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/10/023,929   TIME: 08:42:53

Input Set : N:\Crf3\RULE60\10023929.raw
Output Set: N:\CRF3\02112002\J023929.raw

```
172  145                      150                 155                  160
173  His Asp Arg Ala His Val Ile Lys Lys Ser Lys Asn Lys Lys Thr Gly
174                      165                 170                 175
175  Asp Glu Glu Val Asn Gln Glu Phe Ile Asn Met Asn Glu Ser Asp Ala
176                  180                 185                 190
177  His Ala Phe Asp Glu Glu Trp Gln Ser Glu Val Leu Lys Tyr Lys Pro
178              195                 200                 205
179  Gly Arg His Asn Leu Gly Asn Thr Arg Met Arg Ser Val Gly His Glu
180          210                 215                 220
181  Asn Pro Gly Ser Arg Glu Leu Lys Arg Arg Glu Lys Pro Gln Gln Ser
182  225                      230                 235                  240
183  Pro Ala Ile Glu His Gly Arg Arg Ser Asn Val Leu Gly Asp Lys Leu
184                  245                 250                 255
185  His Ile Lys Gly Ser Ser Val Lys Ser Asn Lys Lys
186              260                 265
188  (2) INFORMATION FOR SEQ ID NO: 4:
190       (i) SEQUENCE CHARACTERISTICS:
191           (A) LENGTH: 1116 base pairs
192           (B) TYPE: nucleic acid
193           (C) STRANDEDNESS: single
194           (D) TOPOLOGY: linear
196     (vii) IMMEDIATE SOURCE:
197           (A) LIBRARY: GenBank
198           (B) CLONE: 1066391
200      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
202  GTTATGTGTT CCCGTCCGTA CTGGAGGCTA GCTCTTGTCG CGGCCGCGGC GAGTTAACAT     60
203  CGTTTTTCCA ATCTGTCCGC GGCTGCCGCC ACCCAAGACA GAGCCAGAAT GTTCAGGATG    120
204  CTGAACAGCA GTTTTGAGGA TGACCCCTTC TTCTCTGAGT CCATTCTTGC ACACCGAGAA    180
205  AATATGCGAC AGATGATAAG AAGTTTTTCT GAACCCTTTG GAAGAGACTT GCTCAGTATC    240
206  TCTGATGGTA GAGGGAGAGC TCATAATCGT AGAGGACATA ATGATGGTGA AGATTCTTTG    300
207  ACTCATACAG ATGTCAGCTC TTTCCAGACC ATGGACCAAA TGGTGTCAAA TATGAGAAAC    360
208  TATATGCAGA AATTAGAAAG AAACTTCGGT CAACTTTCAG TGGATCCAAA TGGACATTCA    420
209  TTTTGTTCTT CCTCAGTTAT GACTTATTCC AAAATAGGAG ATGAACCGCC AAAGGTTTTT    480
210  CAGGCCTCAA CTCAAACTCG TCGAGCTCCA GGAGGAATAA AGGAAACCAG GAAAGCAATG    540
211  AGAGATTCTG ACAGTGGACT AGAAAAAATG GCTATTGGTC ATCATATCCA TGACCGAGCT    600
212  CATGTCATTA AAAAGTCAAA GAACAAGAAG ACTGGAGATG AAGAGGTCAA CCAGGAGTTC    660
213  ATCAATATGA ATGAAAGCGA TGCTCATGCT TTTGATGAGG AGTGGCAAAG TGAGGTTTTG    720
214  AAGTACAAAC CAGGACGACA CAATCTAGGA AACACTAGAA TGAGAAGTGT TGGCCATGAG    780
215  AATCCTGGCT CCCGAGAACT TAAAAGAAGG GAGAAACCTC AACAAAGTCC AGCCATTGAA    840
216  CATGGAAGGA GATCAAATGT TTTGGGGGAC AAACTCCACA TCAAAGGCTC ATCTGTGAAA    900
217  AGCAACAAAA AATAAATAGC CATGCATTTG ATTTGTTTAG TTTTGATTGT TTTAACAGTT    960
218  AGTAATGGTG CTGGGTAATA AGCATAAGAC CAATCTCTTG CTGTTAAATC AGTTCTGTCC   1020
219  TTGGCAACTT TCTTCTGATA TCTGAATGTT CATGAAGGTC CTAGCTTTAT ATTGTCCCTC   1080
220  TTTTAGGAAT AAAATTTTGA TTTTCAACAA AAAAAA                             1116
222  (2) INFORMATION FOR SEQ ID NO: 5:
224       (i) SEQUENCE CHARACTERISTICS:
225           (A) LENGTH: 248 amino acids
226           (B) TYPE: amino acid
227           (C) STRANDEDNESS: single
```

RAW SEQUENCE LISTING   DATE: 02/11/2002
PATENT APPLICATION: US/10/023,929   TIME: 08:42:54

          Input Set : N:\Crf3\RULE60\10023929.raw
          Output Set: N:\CRF3\02112002\J023929.raw

```
228              (D) TOPOLOGY: linear
230      (vii) IMMEDIATE SOURCE:
231              (A) LIBRARY: GenBank
232              (B) CLONE: 1399745
234       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
236   Met Phe Arg Phe Met Arg Asp Val Glu Pro Glu Asp Pro Met Phe Leu
237    1               5                   10                  15
238   Met Asp Pro Phe Ala Ile His Arg Gln His Met Ser Arg Met Leu Ser
239               20                  25                  30
240   Gly Gly Phe Gly Tyr Ser Pro Phe Leu Ser Ile Thr Asp Gly Asn Met
241           35                  40                  45
242   Pro Gly Thr Arg Pro Ala Ser Arg Arg Met Gln Gln Ala Gly Ala Val
243       50                  55                  60
244   Ser Pro Phe Gly Met Leu Gly Met Ser Gly Gly Phe Met Asp Met Phe
245   65                  70                  75                  80
246   Gly Met Met Asn Asp Met Ile Gly Asn Met Glu His Met Thr Ala Gly
247               85                  90                  95
248   Gly Asn Cys Gln Thr Phe Ser Ser Ser Thr Val Ile Ser Tyr Ser Asn
249           100                 105                 110
250   Thr Gly Asp Gly Ala Pro Lys Val Tyr Gln Glu Thr Ser Glu Met Arg
251           115                 120                 125
252   Ser Ala Pro Gly Gly Ile Arg Glu Thr Arg Arg Thr Val Arg Asp Ser
253       130                 135                 140
254   Asp Ser Gly Leu Glu Gln Met Ser Ile Gly His His Ile Arg Asp Arg
255   145                 150                 155                 160
256   Ala His Ile Leu Gln Arg Ser Arg Asn His Arg Thr Gly Asp Gln Glu
257               165                 170                 175
258   Glu Arg Gln Asp Tyr Ile Asn Leu Asp Glu Ser Glu Ala Ala Ala Phe
259           180                 185                 190
260   Asp Asp Glu Trp Arg Arg Glu Thr Ser Arg Phe Arg Gln Gln Arg Pro
261           195                 200                 205
262   Leu Glu Phe Arg Arg Leu Glu Ser Ser Gly Ala Gly Gly Arg Arg Ala
263       210                 215                 220
264   Glu Gly Pro Pro Arg Leu Ala Ile Gln Gly Pro Glu Asp Ser Pro Ser
265   225                 230                 235                 240
266   Arg Gln Ser Arg Arg Tyr Asp Trp
267                   245
269 (2) INFORMATION FOR SEQ ID NO: 6:
271      (i) SEQUENCE CHARACTERISTICS:
272              (A) LENGTH: 1502 base pairs
273              (B) TYPE: nucleic acid
274              (C) STRANDEDNESS: single
275              (D) TOPOLOGY: linear
277      (vii) IMMEDIATE SOURCE:
278              (A) LIBRARY: GenBank
279              (B) CLONE: 1399744
281       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
283   CTCTAAAGGG CAGCTGTGGG AGGAGCGGC GTGGAAGGCC GAGGAGCTCA AGCCCGGACC      60
284   AATCCCCACG TTCCGGGCCG CGACCCTGAC CCTGCAGCGT ACCGGGAAGC GAAACCGGCC     120
```

VERIFICATION SUMMARY   DATE: 02/11/2002
PATENT APPLICATION: US/10/023,929   TIME: 08:42:55

Input Set : N:\Crf3\RULE60\10023929.raw
Output Set: N:\CRF3\02112002\J023929.raw

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:74 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:76 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1